Chapter 5
Hypotheses Test of the Mean

Section 5.2
Conduct the Hypothesis Test

David W. Gerbing

The School of Business
Portland State University

- Conduct the Hypothesis Test
  - Compare Obtained t_m to $t_{\alpha/2}$
  - Compare Obtained p-value to α
  - Application
  - Appendix: Obtain the p-value Directly

5.2a
Compare Obtained t_m to $t_{\alpha/2}$ Criterion

Illustration: t in the Rejection Region

Calculate the t-statistic

- Suppose a claim, the null hypothesis, is that some variable Y has a population mean of 5, $H_0 : \mu = 5$
- Consider an actual sample of data of variable Y for $n = 60$ that yields calculated values of $m = 5.76$ and $s = 2.73$
- Is $m = 5.76$ so much larger than the reference value of 5 that $\mu = 5$ is unreasonable?
- To answer, first estimate the standard error of the mean

$$s_m = \frac{s}{\sqrt{n}} = \frac{2.73}{\sqrt{60}} = 0.352$$

- Then obtain the observed standardized distance of m from μ_0

$$t_m = \frac{m - \mu_0}{s_m} = \frac{5.76 - 5}{0.352} = 2.16$$

Illustration: t in the Rejection Region

- *Criterion*: Start with the theoretical t-distribution and $\alpha = 0.05$ criterion without any reference to the data
- *From the data*: $t_m = 2.16$, so the sample mean is 2.16 estimated standard errors above the hypothesized mean
- *Evaluate*: compare obtained $t_m = 2.16$ with upper tail cutoff $t_{.025;df=59} = 2.00$

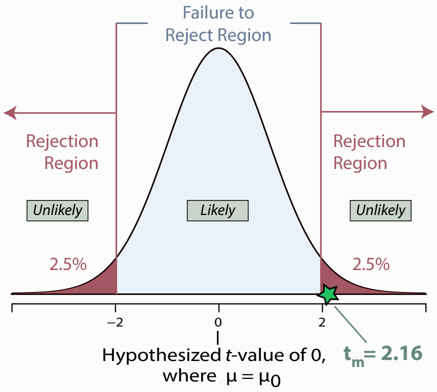


Illustration: t in the Rejection Region

- For $df = 59$, $t_m = 2.16 > t_{.025} = 2.00$, so t_m lies in the upper tail rejection region
- IF $H_0 : \mu = \mu_0$ is true, THEN obtained outcome of t_m is unlikely, so reject $H_0 : \mu = 5$
- Conclude that $\mu = 5$ is unreasonable

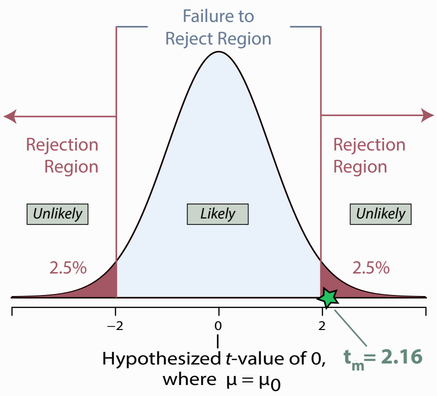


Illustration: t NOT in the Rejection Region

Calculate the t-statistic

- Suppose the claim is the null hypothesis for some variable Y is that its population mean is 5, $H_0 : \mu = 5$
- Consider an actual sample of data of variable Y for $n = 60$ that yields calculated values of $m = 5.37$ and $s = 2.73$
- Is $m = 5.37$ so much larger than the reference value of 5 that $\mu = 5$ is not reasonable?
- First estimate the standard error of the mean

$$s_m = \frac{s}{\sqrt{n}} = \frac{2.73}{\sqrt{60}} = 0.352$$

- Then obtain the observed standardized distance of m from μ_0

$$t_m = \frac{m - \mu_0}{s_m} = \frac{5.37 - 5}{0.352} = 1.05$$

Illustration: t NOT in the Rejection Region

- *Criterion*: Start with the theoretical t-distribution and $\alpha = 0.05$ criterion without any reference to the data
- *From the data*: $t_m = 1.05$, so the sample mean is 1.05 estimated standard errors above the hypothesized mean
- *Evaluate*: compare obtained $t_m = 1.05$ with upper tail cutoff $t_{.025;df=59} = 2.00$

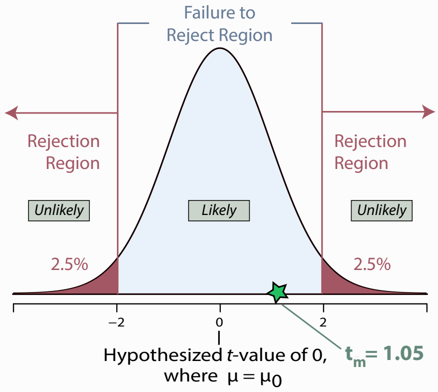
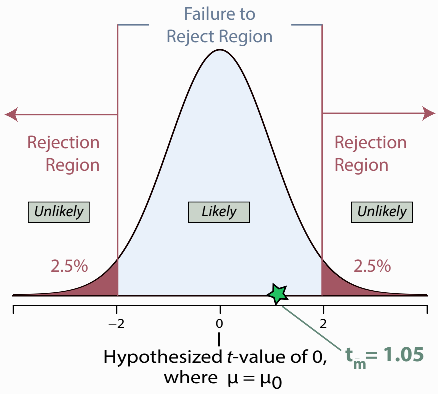


Illustration: t NOT in the Rejection Region

- For $df = 59$, $t_m = 1.05 < t_{.025} = 2.00$, so t_m does NOT lie in a rejection region
- IF $H_0 : \mu = \mu_0$ is true, THEN t_m is reasonably likely, so do *not* reject $H_0 : \mu = 5$
- Conclude: No difference from $\mu = 5$ detected
- Do *not* conclude that $\mu = 5$ is the correct value



5.2b
Compare Obtained p-value to α Criterion

Assess How Close m is to μ_0 with Probability

IF μ_0 is true, THEN how probable is the sample value t_m?

- p-**value**: Given a true null hypothesis, the probability of obtaining a value as far from or farther from the hypothesized value the value of the obtained sample mean
- Usually either a large positive or a large negative deviation from the hypothesized value is of interest, so the p-value usually assesses the probability of a value in either tail
- IF $\mu = \mu_0$, THEN
  - A low p-value indicates an *unlikely* sample mean, so t_m is in the rejection region, which means m is far from μ_0
  - A high p-value indicates a *likely* sample mean, so t_m is not in the rejection region, and m is reasonably close to μ_0
- Prefer more convenient probability expression of the p-value because probabilities are more generally understood than a t-value, and both provide the identical result

The p-value Approach to the Hypothesis Test

Specify what is meant by “likely” and “unlikely”

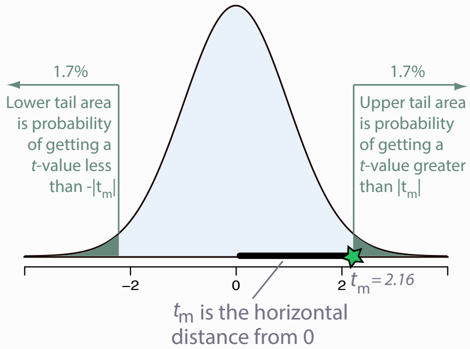
- Usually the definition of the threshold for “unlikely”, called alpha or α, is set at $\alpha = 0.05$
- Compare p-value to α to see if obtained t_m is unlikely
- The approach of comparing t_m to $t_{\alpha/2}$ is more “old school” now that obtaining the p-value is more feasible with computers

Statistical Decision

- p-value $> \alpha$, then fail to detect a difference from μ_0
- p-value $< \alpha$, then t_m is in the rejection region, so reject μ_0

Obtain p-value: Reject H_0

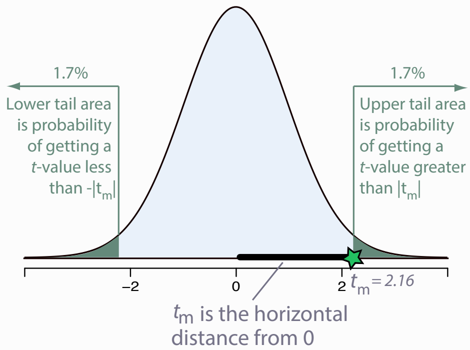
- Suppose $t_m = 2.16$
- t_m: Distance of m from μ_0 in terms of s_m
- For any one obtained $t_m > 0$, there is a corresponding upper-tail area, here 0.017
- To account for deviations in either direction, also consider corresponding lower-tail area
- Total area in *both* tails: p-value $= (2)(.017) =$ .034



Compare p-value with $\alpha = .05$: Reject H_0

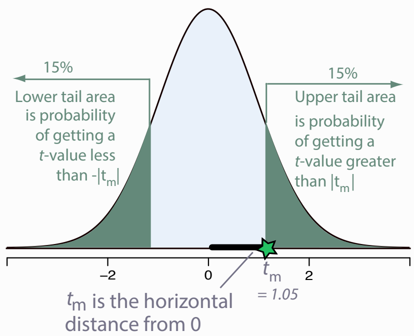
- p-value $= .034 < \alpha = .05$

  IF μ_0 is true, THEN a m that far from μ_0 in either direction, or farther, has a low probability of occurring, only 0.034
- A low p-value means that, assuming μ_0 is true, the sample result is UNLIKELY
- So reject the null hypothesis and conclude that apparently $\mu \neq \mu_0$



Obtain p-value: Do *not* Reject H_0

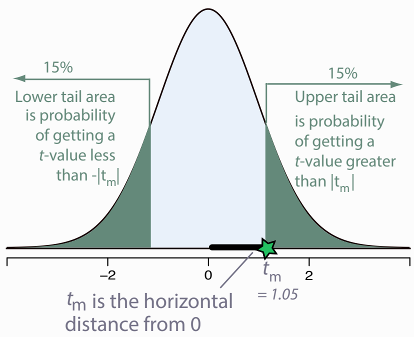
- Suppose $t_m = 1.05$
- t_m: Distance of m from μ_0 in terms of s_m
- For any one $t_m > 0$, there is a corresponding upper-tail area, here 0.15
- To account for deviations in either direction, also consider corresponding lower-tail area
- Total area in *both* tails: p-value $= (2)(.15) = .30$



Compare *p*-value with $\alpha = .05$: Do *not* Reject H_0

- *p*-value $= .30 > \alpha = .05$

  IF μ_0 is true, THEN a *m* that far from μ_0, or farther, in either direction has a reasonable probability of occurring, 0.30
- A high *p*-value means that the sample result is CONSISTENT with the assumption of μ_0
- So do *not* reject the null hypothesis and conclude that no difference detected from μ_0



Interpretation of the Hypothesis Test

The statistical decision: Reject or Do not reject

- Fail to Reject the Null when ...
  - *p*-value $> \alpha$: Sample mean is close to μ_0
  - **Conclude**: No difference from μ_0 detected
  - Conclusion of no difference is NOT that μ_0 is true
- Reject the Null when ...
  - *p*-value $< \alpha$: Sample mean far from μ_0, in rejection region
  - **Conclude**: A difference from the specified null hypothesize value of the mean is detected
  - For purposes of making a decision, simply detecting a difference does not provide any actionable information
  - Instead, what is important is the direction and extent of the difference, best evaluated with a confidence interval to follow-up the hypothesis test

p-value is a Conditional Probability

Distinguish between what we want and what we have

- **Conditional probability**: The probability of an event only when some other condition is true
- Notation: $P(A \mid B)$, read as the probability of event A given condition B
- A *p*-value is a conditional probability: *p*-value $= P(data \mid H_0)$
- The *p*-value is the probability of obtaining the data, and corresponding sample mean, *given* that the value of the mean specified by the null hypothesis is true
- Unfortunately, we actually want to know the probability that the null hypothesis is true given the data, $P(H_0 \mid data)$
- ☹ What we want is not what we get
- And knowing one conditional probability does not inform us as to the value of the other: $P(data \mid H_0) \neq P(H_0 \mid data)$

Hypothesis Tests Never Provide the Probability of the Null

Conclusion of a hypothesis test is only qualitative

- **Key Concept**: Despite the precise calculation of a *p*-value, an hypothesis test provides a *qualitative result*, reasonable or unreasonable
- IF μ_0 is assumed true,
  - THEN when the sample result is unlikely, *p*-value $< .05$, conclude μ_0 is *unreasonable*
  - THEN when the sample result is likely, *p*-value $> .05$, conclude μ_0 is *reasonable*
- Computers can precisely calculate the *p*-value to as many decimal digits as desired
- The problem is that the probability of what we really want, the probability that μ_0 is the true value of μ, is not known, only qualitatively described as reasonable or unreasonable

R: One-sample Hypothesis Test

Analysis provided by `lessR` function `ttest`, or `tt`

- Analysis of variable Y from data: `> ttest(Y, mu=value)`
- Ex: To test that the mean weight of cereal in the cereal boxes is the specified 350g for the variable Weight

```
> ttest(Weight, mu=350)
```

- The output of `ttest` includes the relevant *p*-value
- The brief version, `tt_brief()`, provides just the descriptive statistics, the hypothesis test and confidence interval
- The output provides a graph of the smoothed histogram (densities) of the variable of interest
- Analysis of Y from sample size, mean and standard deviation:

```
> ttest(n=value, m=value, s=value, mu=value)
```


5.2c
Application

Assess Average Weight of Boxed Cereal

Claim: Advertised weight in each box is 350 g

- Inherent process variability ensures that the contents of each box is typically more or less than 350 g
- But is there a systematic over- or under-fill as the cereal boxes are filled?
- Decision: On the production line, do we adjust the dial – up or down – to increase or decrease the average amount of cereal placed in each cereal box?

- Null Hypothesis is $H_0 : \mu = 350$g
- Alternative Hypothesis is $H_1 : \mu \neq 350$g

- Data: Randomly sample 25 cereal boxes and measure the weight of the cereal in each box

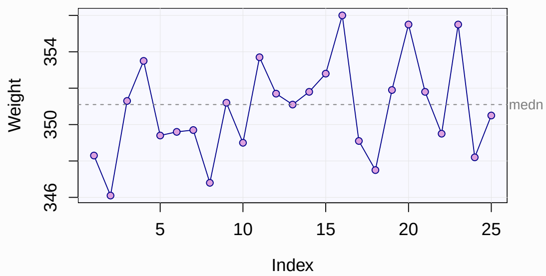
`http://lessRstats.com/data/boxweight.csv`

Weight
348.3
346.1
351.3
353.5
349.4
349.6
349.7
346.8
351.2
349.0
353.7
351.7
351.1
351.8
352.8
356.0
349.1
347.5
351.9
355.5
351.8
349.5
355.5
348.2
350.5

Stable Process Assumption of Y

Only meaningful to estimate a stable μ

- Are all data values of Y, Weight, from the same process?
- To answer, plot the run chart: `> LineChart(Weight)`

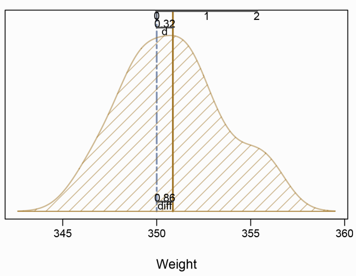


- Output appears to be random variation, so process stable
- Conclude: There is a stable μ to estimate

Normality Assumption of Sample Mean m

t-distribution probabilities only valid for normal m

- The distribution over repeated samples of m for n close to 30 is approximately normal *unless* Y is skewed
- To verify, examine the graph from the `ttest` function, a smoothed graph (the densities) of the distribution of Y



- The sample data are approximately normal so m, the mean of Weight, appears normal across multiple, hypothetical samples

Excel Template: Descriptive Statistics

	Description	Name	Value	Formula
INPUT: DESCRIPTIVE STATISTICS	count of data	n	25	COUNT(data)
	mean of data	mean	350.86	AVERAGE(data)
	standard dev of data	stdev	2.66	STDEV(data)

Traditional Notation:
$n = 25$
$m = 350.86$
$s = 2.66$

Same three basic descriptive statistics for calculating the confidence interval are also the basis for the hypothesis test

Basic Question:
Is sample mean of 350.860 g close to hypothesized mean of 350 g?

Excel Template: Standard Error

	Description	Name	Value	Formula
INPUT: DESCRIPTIVE STATISTICS	count of data	n	25	COUNT(data)
	mean of data	mean	350.860	AVERAGE(data)
	standard dev of data	stdev	2.660	STDEV(data)
	std error of mean	sterr	0.532	stdev/SQRT(n)

Traditional Notation:

$$s_m = \frac{s}{\sqrt{n}} = \frac{2.66}{\sqrt{25}} = 0.532$$

Same estimated standard error for calculating the confidence interval is also the basis for the hypothesis test

How many estimated standard errors separate 350.860 g from hypothesized mean of 350 g?

Excel Template: t-value

	Description	Name	Value	Formula
INPUT: DESCRIPTIVE STATISTICS	count of data	n	25	COUNT(data)
	mean of data	mean	350.860	AVERAGE(data)
	standard dev of data	stdev	2.660	STDEV(data)
	std error of mean	sterr	0.532	stdev/SQRT(n)
HYPOTHESIS TEST	hypothesized value	mu0	350	
	difference from null	diff	.86	mean-mu0
	t-value (distance)	t	1.617	diff/stderr
	p-value	p-value	0.119	TDIST(ABS(t),n-1,2)

Now specify the value of the hypothesized value, μ_0, and then obtain the obtained t-value, followed by its corresponding p-value

$$t_m = \frac{m - \mu_0}{s_m} = \frac{350.86 - 350}{0.532} = \frac{0.860}{0.532} = 1.616$$

which yields a p-value $= 0.119$

R: Inference of the Mean, HT and CI, with lessR ttest

```
> tt_brief(Weight, mu=350)
```

```
Weight: n = 25,   mean = 350.86,   sd = 2.66

Hypothesized Value H0: mu = 350
Hypothesis Test of Mean:
    t-value = 1.62,  df = 24,  p-value = 0.119

95% Confidence Interval for Mean:  349.76 to 351.96
```

Or, run the analysis directly from the summary statistics

```
> tt_brief(n=25, m=350.86, s=2.66, mu=350)
```

To obtain more information, run the full version of `ttest`

```
> ttest(Weight, mu=350)
```


Result

Result of the Analysis

- Statistical Result: $t_m = \frac{m - \mu_0}{s_m} = 1.616$
  - The sample mean $m = 350.86$g is 1.616 estimated standard errors above the hypothesized mean μ_0 which is $\mu = 350$g
  - Deviations from the hypothesized value of 350g in either direction are of interest, so is the distance of 1.616 standard errors or more from $\mu = 350$g in *either* direction likely?
- Statistical Result: At $t_m = 1.616$, $df = 24$ the p-value $\approx 12\%$
  - For a true mean of $\mu = 350$g, the probability of obtaining a value of m at or more than 0.86 grams on either side of 350g is 12%
  - 12% does not indicate a particularly rare event

Conclusion

Implications of the statistical decision

- **Statistical Decision**:

  $$p - \text{value} = 0.119 \quad > \quad \alpha = .05$$

  so do *not* reject the null hypothesized value of 350 g, the sample mean of 350.86 is close to 350
- **Interpretation**: No difference in the average weight of the contents of the cereal boxes detected from 350 g
- **Caveat**: Not rejecting the null hypothesis of $\mu = 350$ g does not imply the true mean is 350 g
- **Managerial Decision**: Do *not* adjust the dial that controls how much cereal on average is placed in each cereal box
- **Follow-up Analysis**: When the null hypothesis is not rejected, something called power analysis should be conducted, explained later

Appendix
Obtain the *p*-value Directly

Obtain the *p*-value

Technology moves on

- In the "old" days, accurate *p*-values were hard to obtain because a table of *t*-values does not typically provide enough probabilities for each *t*-distribution corresponding to a specific degrees of freedom
- Today *p*-values are automatically provided by the computer when conducting an hypothesis test, with applications such as R and Excel
- The *p*-value section of computer output of the hypothesized test is usually considered the most crucial part of the displayed analysis
- Also can obtain a *p*-value with programs such as R and Excel apart from a specific data analysis, analogous to the "old school" method of looking up a value in a printed *t*-table, shown in the Appendix

R: Manually Obtain a *p*-value

The R `pt` function

- Use the probability *t*-distribution function, `pt`, which by default provides the cumulative probability, the "lower tail" of the distribution, appropriate for obtained negative values of t_m
- For the usual test that investigates possibilities in both tails, multiply the obtained tail probability by two to get the probability of a deviation as large or larger in *either* tail
- Ex: Obtained $t_m = -1.616$, df=24

```
> 2*pt(-1.616, df=24)
```
returns 0.119

- Set `lower.tail=FALSE` to get the upper tail probability, which is usually desired for positive values of t_m
- Ex: Obtained $t_m = 1.616$, df=24

```
> 2*pt(1.616, df=24, lower.tail=FALSE)
```
returns 0.119

Excel: Manually Obtain a p-value

Use the TDIST function

- Use the TDISTribution function to get the probability of one or both tails of the specified t-distribution

```
=TDIST(ABS(t-value), df, number of tails)
```

- Only positive t-statistics can be specified, so first call the ABSsolute value function or only enter a positive t-value
- Ex: for $t_m = 1.616, df = 24$

`=TDIST(ABS(1.616), 24, 2)` returns 0.119

- Excel has no one-sample t-test function
- An Excel template for this t-test which uses the TDIST function is provided in the following application

Index Subtract 2 from each listed value to get the Slide

- The End